International Research Journal of Innovative Engineering (IRJIE)

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 8, Issue 4, July- August 2024|

DOI:10.15680/IRJIE.2024.0804004

Modern Business Intelligence: Big Data Analytics and Artificial Intelligence for Creating the Data-Driven Value

Maheshkumar Mole

Motorola Solutions, Inc., USA

Email: maheshmole@gmail.com

https://orcid.org/0009-0001-5966-9265

ABSTRACT: The integration of Artificial Intelligence into Business Intelligence and cloud-native big data platforms represents a paradigm shift in how organizations process, analyze, and derive strategic value from vast data repositories. This article provides a comprehensive, technically rigorous examination of AI-powered Business Intelligence, charting the transition from traditional descriptive analytics to predictive and prescriptive intelligent systems. Grounded in foundational literature across automated machine learning, explainable artificial intelligence, real-time anomaly detection, and natural language processing, this analysis deconstructs the algorithms powering modern analytical features. The architectural underpinnings of cloud-native data platforms utilizing Medallion architectures and automated pipelines are thoroughly explored. Furthermore, the mathematical formulations of core capabilities—such as the Isolation Forest algorithm for multidimensional anomaly detection, SHapley Additive exPlanations for model interpretability, and Auto-Regressive Integrated Moving Average models for time-series forecasting—are detailed to explicate their successful operationalization in enterprise tools like Power BI and Tableau. The analysis concludes by quantifying the economic impact of these technologies, demonstrating how AI-augmented data platforms facilitate smart governance, democratize data access through cross-domain semantic parsing, and deliver substantial organizational returns on investment.

KEYWORDS: Artificial Intelligence, Augmented Analytics, Automated Machine Learning, Business Intelligence, Cloud-Native Architecture, Explainable AI, Isolation Forest, Semantic Parsing, Time Series Forecasting.

I. INTRODUCTION

BI's transformation into an AI-driven discipline stems from decades of advances in data processing, storage, and algorithms. Contextualizing this within the progression of analytics illuminates the current state of AI-powered BI.

A. The Evolution from BI & A 1.0 to 3.0

The evolution of Business Intelligence and Analytics (BI&A) into three distinct maturity phases.

BI&A 1.0 focused on structured data in relational database management systems (RDBMS). It relied on descriptive analytics, such as OLAP, standard batch reporting, and basic data mining, answering "what happened" through intensive manual effort.

BI&A 2.0 leveraged the Internet and Web 2.0 to analyze unstructured web data, text, customer opinions, and social media. BI&A 3.0 expanded into mobile devices, sensors, and IoT, spurring the growth of "Big Data".

Modern AI-powered BI operationalizes BI&A 3.0, resolving human cognitive limits as data volume, velocity, and variety grow. Integrating machine learning into BI workflows enables automated insight extraction from exabyte-scale datasets.

International Research Journal of Innovative Engineering (IRJIE)

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 8, Issue 4, July- August 2024|

DOI:10.15680/IRJIE.2024.0804004



Figure 1: Historical Evolution of Business Analytics (BI&A 1.0 to 3.0)

II. ARCHITECTURAL FOUNDATIONS: CLOUD-NATIVE BIG DATA PLATFORMS

The sophisticated AI features observed in modern BI applications are strictly dependent on the underlying data architecture. The integration of AI into BI cannot occur effectively in isolated desktop environments or fragmented data silos; it requires a highly scalable, distributed, and strictly governed cloud-native data foundation.

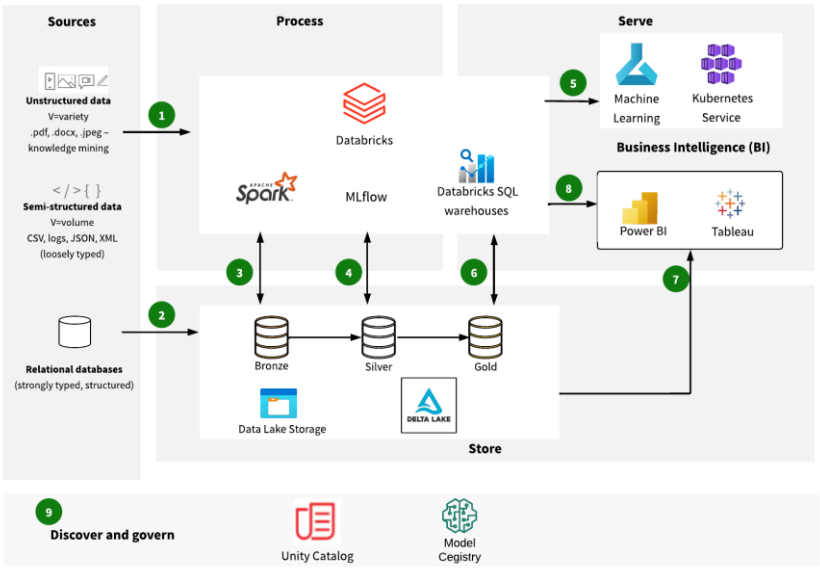


Figure 2: Architecture and Design for data platform and Business Intelligence

A. The Medallion Data Architecture

Cloud-native big data platforms (e.g., Databricks, Azure Synapse, Google BigQuery, Snowflake) use the Medallion Architecture. This design pattern concurrently supports traditional BI and machine learning models without resource contention.

International Research Journal of Innovative Engineering (IRJIE)

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 8, Issue 4, July- August 2024|

DOI:10.15680/IRJIE.2024.0804004

The Medallion Architecture structures data into three progressive layers, continuously increasing quality and validation:

1. **Bronze Layer (Raw Data Ingestion):** Incalculably lands raw, unprocessed data (JSON, Parquet, CSV) from various sources (databases, IoT, APIs, logs) to maintain an immutable historical archive.
2. **Silver Layer (Cleansed and Conformed):** Cleans, standardizes, and joins data, addressing anomalies and nulls. It serves as a reliable base for data science and machine learning pipelines.
3. **Gold Layer (Curated for Analytics):** Stores highly refined, aggregated data modeled (e.g., star/snowflake schemas) specifically for direct, low-latency querying by BI reporting tools.

Layer	Function	Data State	Primary Consumers
Bronze	Raw ingestion & immutable archiving	Unprocessed (JSON, CSV)	Data Engineers, Pipelines
Silver	Cleansing, conforming, deduplication	Standardized, deduplicated	Data Scientists, ML Models, Analysts
Gold	Business aggregation & dimensional modeling	Structured, optimized star schema	BI Tools (Power BI, Tableau)

B. Integration of Machine Learning Operations (MLOps)

Cloud-native data platforms integrate AI capabilities and lifecycle management. Frameworks like MLflow manage the ML lifecycle, allowing data scientists to track parameters, package reproducible runs, and deploy predictive models across the enterprise.

These architectures use Unity Catalogs and Model Registries for data governance and access control, while supporting distributed training with frameworks like TensorFlow, PyTorch, and Keras. Co-locating compute power with Big Data storage removes data movement bottlenecks, enabling real-time model inference and seamless BI integration so analysts can query data via PySpark, Koalas, or SQL from a unified interface.

III. THE ROLE OF AUTOMATED MACHINE LEARNING (AUTOML)

A defining characteristic of AI-powered BI is the abstraction of highly complex data science tasks away from the end-user. This democratization is achieved primarily through Automated Machine Learning (AutoML).

A. The Foundations of AutoML

AutoML automates machine learning for real-world applications. Hutter, Kotthoff, and Vanschoren (2019) note that rapid growth and a scarcity of data scientists drove demand for accessible methods. AutoML addresses this by automating data preprocessing, feature selection, model selection, hyperparameter optimization (HPO), and neural architecture search (NAS).

It resolves the Combined Algorithm Selection and Hyperparameter Optimization (CASH) problem by searching algorithm and hyperparameter combinations for optimal performance.

Techniques like Bayesian Optimization, Evolutionary Algorithms, Genetic Programming, and Meta-learning navigate these search spaces. For example, TPOT uses genetic programming to optimize pipelines with minimal user input.

B. AutoML Operationalized in BI Tools

Modern BI platforms leverage AutoML behind the presentation layer. User interactions trigger automated pipelines that preprocess data, train models, and tune hyperparameters in real time.

International Research Journal of Innovative Engineering (IRJIE)

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 8, Issue 4, July- August 2024|

DOI:10.15680/IRJIE.2024.0804004

AutoML Function	Traditional Data Science Approach	AI-Powered BI Approach
Feature Selection	Manual statistical testing and expert curation	Automated extraction and ranking based on variance and correlation
Model Selection	Iterative testing of algorithms (e.g., SVM vs. Random Forest)	Parallel algorithm execution via the CASH framework
Hyperparameter Tuning	Manual grid or random search	Meta-learning and Bayesian optimization for rapid parameter convergence
Insight Generation	Manual Python/R coding for visualizations	Dashboard-integrated narrative summaries and auto-generated charts

For instance, engaging "Quick Insights" in a BI platform deploys machine learning algorithms to auto-detect patterns and outliers. The AutoML engine searches across linear models, decision trees, and clustering algorithms, rendering significant correlations directly as visuals for AI-guided discovery.

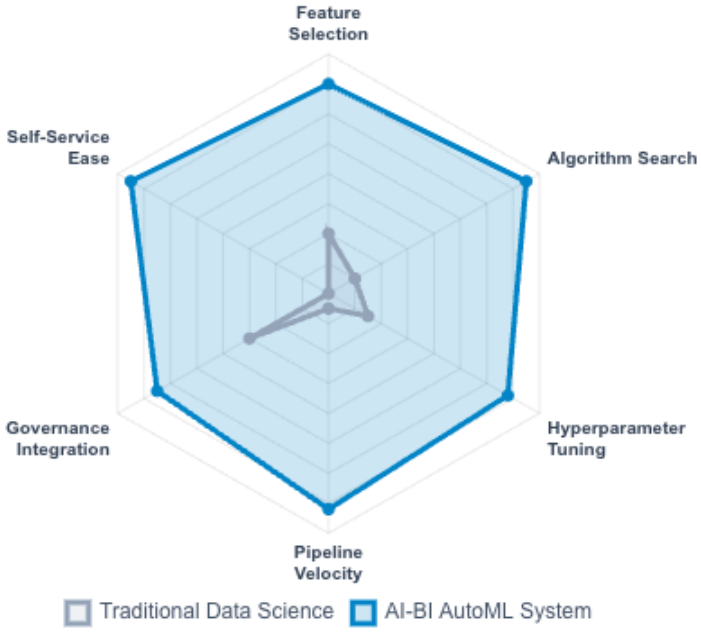


Figure 3: Traditional Modeling vs. AI-BI AutoML Capability

IV. NATURAL LANGUAGE INTERFACES AND CROSS-DOMAIN SEMANTIC PARSING

One of the most visible and commercially impactful integrations of AI into BI is the ability for users to interrogate their enterprise data using natural language. This capability drastically lowers the barrier to entry for data analysis, enabling non-technical users to extract complex information simply by asking questions in plain language.

A. The Evolution of Natural Language Interfaces to Databases (NLIDB)

The quest to query databases using natural language has a long and complex history in computer science, typically categorized under the domain of Natural Language Interfaces to Databases (NLIDB). Historically, end-users who did

International Research Journal of Innovative Engineering (IRJIE)

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 8, Issue 4, July- August 2024|

DOI:10.15680/IRJIE.2024.0804004

not proficiently speak formal query languages like SQL or SPARQL were effectively excluded from exploring database systems. Early NLIDB systems relied on rigid, hand-crafted rules and narrow, domain-specific ontologies, making them exceptionally brittle. If a user asked a question using a synonym not explicitly programmed into the ontology, the system would fail. Consequently, these early systems were nearly impossible to scale across different enterprise domains.

The breakthrough for modern BI tools came with advancements in deep learning, specifically in the area of large-scale semantic parsing. Semantic parsing is the task of translating natural language utterances into executable formal representations, such as SQL queries.

B. The Spider Dataset and Cross-Domain Generalization

Yu et al. (2018) introduced Spider, a benchmark dataset for cross-domain semantic parsing and text-to-SQL tasks. Earlier models were limited to specific individual databases and failed to generalize.

Spider evaluated models on unseen, complex database schemas featuring multiple tables, foreign keys, and advanced SQL operations. Initial sequence-to-sequence models managed only 12.4% exact matching accuracy on database splits, highlighting severe performance limits.

To address this, modern systems employ advanced schema linking and graph neural networks to capture relational database structures before generating SQL.

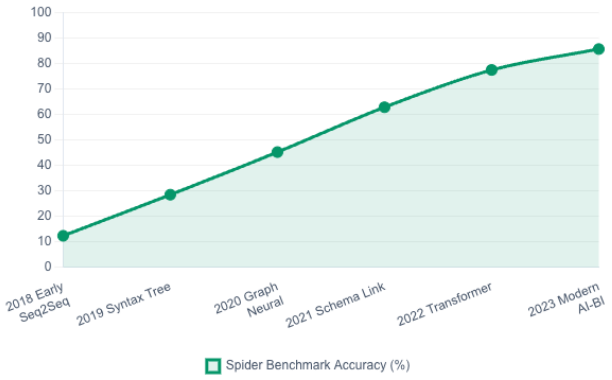


Figure 4: Spider Benchmark: Text-to-SQL Accuracy Evolution

C. Conversational BI and Generative Assistants

Leading BI platforms use advanced text-to-SQL and semantic parsing to power tools like Tableau's "Ask Data".
When queried (e.g., "What were the sales in 2022 by product category?"), the NLP engine parses the query against the Semantic Data Model, generates the required DAX or SQL, executes it, and displays the response using the best visualization type.

Integrated foundational language models allow these assistants to understand nuanced user intent, summarize reports, explain trends, and refine charts through follow-up prompts—acting as an on-demand "AI-Powered Business Analyst."

V. PREDICTIVE ANALYTICS: MATHEMATICAL FOUNDATIONS OF FORECASTING

While descriptive analytics explains what has historically occurred, predictive analytics leverages historical data to forecast what will happen in the future. BI platforms have deeply integrated advanced time-series forecasting techniques, allowing organizations to visualize projected data, confidence intervals, and trends to radically enhance enterprise planning processes.

A. Exponential Smoothing (ETS) Formulations

Exponential Smoothing (ETS) is a foundational forecasting method for modeling time-series data, identifying patterns, and mapping dependencies over time. Unlike simple moving averages that assign equal weight to all past observations

International Research Journal of Innovative Engineering (IRJIE)

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 8, Issue 4, July- August 2024|

DOI:10.15680/IRJIE.2024.0804004

regardless of their age, exponential smoothing assigns exponentially decreasing weights as the observations get older, making it highly responsive to recent changes in the data.

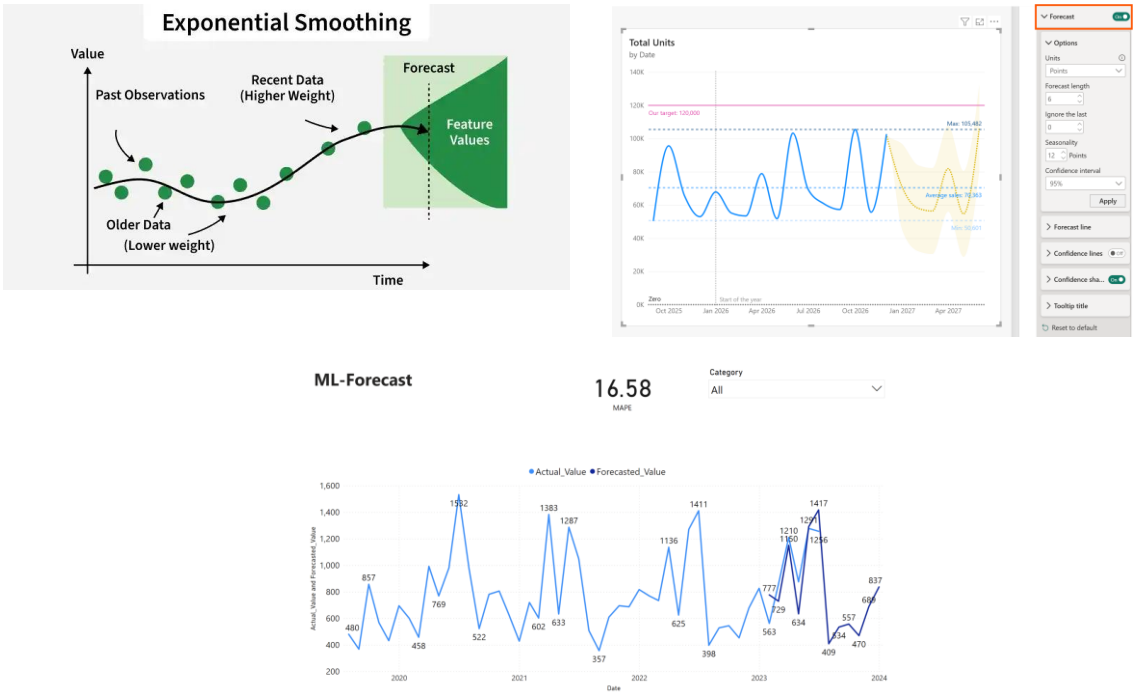


Figure 5: Forecasting sales and demand

The mathematical formulation for a simple exponential smoothing model is given by the recursive equation:

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha)\hat{y}_t$$

Where:

- $\hat{y}_{t+1}$ represents the forecasted value for the subsequent period.
- y_t is the actual observed value at period t.
- $\hat{y}_t$ is the forecasted value for period t.
- α is the smoothing parameter, bounded strictly between 0 and 1 ($0 < \alpha < 1$).

For datasets exhibiting persistent trends, Holt's linear trend method is utilized, which introduces a second smoothing equation to account for the trend component (b_t):

$$\hat{y}_{t+h} = \ell_t + hb_t$$

$$\ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} + b_{t-1})$$
$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$$

Modern BI tools automate the selection of the precise ETS variant (such as adding seasonal components via the Holt-Winters method) by systematically optimizing the Akaike Information Criterion (AIC) or the Bayesian Information

International Research Journal of Innovative Engineering (IRJIE)

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 8, Issue 4, July- August 2024|

DOI:10.15680/IRJIE.2024.0804004

Criterion (BIC). This ensures the selection of the model that best fits the historical data without falling victim to overfitting.

B. Auto-Regressive Integrated Moving Average (ARIMA)

For more complex datasets containing non-stationary elements, BI tools deploy ARIMA models. ARIMA is a robust statistical model that combines autoregression (AR), differencing (I) to achieve stationarity, and moving averages (MA).

An ARIMA(p, d, q) model is mathematically expressed as:

$$\left(1 - \sum_{i=1}^{p} \phi_i L^i\right) (1 - L)^d y_t = \left(1 + \sum_{j=1}^{q} \theta_j L^j\right) \epsilon_t$$

Where:

- L acts as the lag operator.
- p denotes the order of the autoregressive part.
- d represents the degree of first differencing required to make the series stationary.
- q is the order of the moving average part.
- ϕ_i and θ_j are the learned parameters of the model.
- ϵ_t represents the error term or white noise.

Big data platforms use dynamic updating (e.g., expanding-window frameworks) to continuously adjust ARIMA orders, enabling real-time model adaptation without manual data scientist intervention. Embedding these models into BI tools allows users to project trends easily. For instance, clicking a "Forecast" button on a dashboard triggers underlying ETS or ARIMA routines to display future trends and valid confidence intervals automatically.

VI. REAL-TIME ANOMALY AND FRAUD DETECTION

Identifying and preventing anomalous activities, system failures, or fraudulent transactions in real-time is a critical capability enabled by AI in modern BI systems. Traditional rule-based logic is inherently brittle and fails to scale; conversely, machine learning anomaly detection models adapt dynamically to complex, high-dimensional data structures.

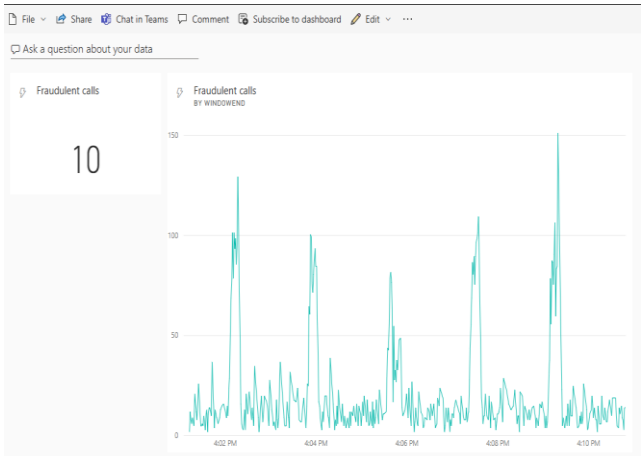


Figure 6: Real-time fraud detection

International Research Journal of Innovative Engineering (IRJIE)

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 8, Issue 4, July- August 2024|

DOI:10.15680/IRJIE.2024.0804004

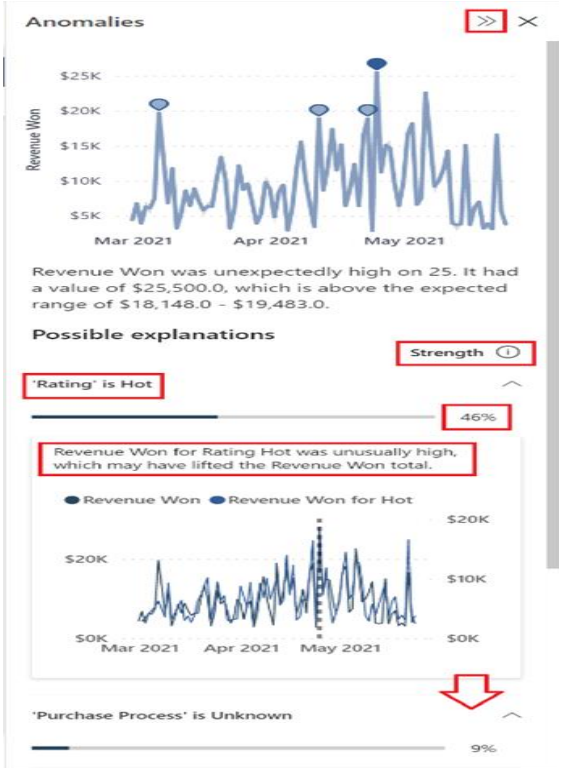


Figure 7: Anomaly Detection

A. Mathematical Formulation of the Anomaly Score

The path length $h(x)$ of a data point x is measured by the number of edges x traverses in an iTree from the root node until the traversal is terminated at an external node. Since iTrees share the exact same underlying structure as Binary Search Trees (BST), the average path length of an unsuccessful search in a BST can be utilized to normalize the path lengths.

For a dataset of n instances, the average expected path length $c(n)$ is calculated as:

$$c(n) = 2H(n-1) - \frac{2(n-1)}{n}$$

Where $H(i)$ is the harmonic number, which can be mathematically estimated by $\ln(i) + 0.5772156649$ (Euler's constant).

The Isolation Forest algorithm aggregates the path lengths across a vast collection of t trees (the forest) to find the expected path length $E(h(x))$ for a given point. The final, normalized anomaly score $s(x,n)$ is then defined as:

$$s(x,n) = 2^{-\frac{E(h(x))}{c(n)}}$$

This scoring formulation provides strict, interpretable boundaries that make it exceptionally useful for automated BI dashboards:

- If $E(h(x)) \to 0$, then $s \to 1$. An anomaly score very close to 1 strongly indicates a severe anomaly.
- If $E(h(x)) \to c(n)$, then $s \to 0.5$. An anomaly score around 0.5 indicates a perfectly normal,

International Research Journal of Innovative Engineering (IRJIE)

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 8, Issue 4, July- August 2024|

DOI:10.15680/IRJIE.2024.0804004

expected data point.

- If $E(h(x)) \to n-1$, then $s \to 0$. The instance is deeply embedded in a dense cluster of normal points and is exceptionally safe.

VII. EXPLAINABLE AI (XAI) IN AUGMENTED ANALYTICS

As BI systems transition from simply displaying historical data to making AI-driven predictions and uncovering automated patterns, the fundamental issue of trust becomes paramount. If business users do not trust a model, or cannot understand its underlying reasoning, they will simply refuse to utilize its outputs for high-stakes decision-making. This critical challenge is addressed through the integration of Explainable AI (XAI) within augmented analytics.

A. The Crisis of Trust and Local Interpretable Models

Ribeiro et al. (2016) highlighted the need for interpretability via LIME, distinguishing trust in individual predictions from trust in a model's overall behavior.

To demystify opaque black-box models like deep networks, LIME explains predictions by locally approximating the complex model with a sparse linear regression model trained on perturbed input data.

B. A Unified Framework: SHapley Additive exPlanations (SHAP)

While LIME provided a strong mechanism for local explanation, the field required a more unified, mathematically rigorous framework to ensure consistency. This was achieved by Lundberg and Lee (2017) with the introduction of SHAP (SHapley Additive exPlanations). SHAP unifies six prior explanation methods (including LIME and DeepLIFT) under a single, overarching class defined as "additive feature attribution methods".

SHAP is deeply grounded in cooperative game theory, specifically leveraging Shapley values introduced in 1953, which provide a method for fairly distributing a total surplus among a coalition of players. In the context of machine learning explainability, the "game" is the prediction task, and the "players" are the input features.

Lundberg and Lee defined an explanation model $g(z')$ as a linear function of binary variables:

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i z_i'$$

Where $z' \in \{0,1\}^M$ represents a simplified feature vector (indicating whether a feature is present or absent), M is the total number of simplified input features, and $\phi_i \in \mathbb{R}$ is the feature attribution (the SHAP value) for feature i.

SHAP is mathematically proven to be the *only* additive feature attribution method that simultaneously satisfies three highly desirable theoretical properties:

1. **Local Accuracy:** The explanation model must exactly match the output of the original model for the specific input being explained. Mathematically, $f(x) = g(x') = \phi_0 + \sum_{i=1}^{M} \phi_i x_i'$.
2. **Missingness:** If a feature is missing in the original input ($x_i' = 0$), it is mathematically constrained to have absolutely no attributed impact ($\phi_i = 0$).
3. **Consistency:** If the underlying model changes such that the marginal contribution of a specific feature value increases or stays the same (regardless of other features), the SHAP value for that feature must never decrease.

By permuting through the vast feature space and computing the expected marginal contribution of each feature across all possible subsets, SHAP provides a highly accurate, theoretically sound explanation that perfectly aligns with human intuition.

International Research Journal of Innovative Engineering (IRJIE)

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 8, Issue 4, July- August 2024|

DOI:10.15680/IRJIE.2024.0804004

C. XAI Application in BI: Key Influencers and Decomposition

Modern BI platforms leverage XAI frameworks to simplify complex AI predictions into visual narratives for business users.

Tableau Explain Data & Einstein Discovery: Tableau's "Explain Data" automatically clarifies data point variances. When selecting an outlier, background regression models and XAI algorithms calculate dimensional contributions to present findings in plain language.

Power BI Key Influencers: Using LIME and SHAP principles, this visual ranks key drivers for target metrics (e.g., factors causing a "Won" sales status) and displays their marginal impact without requiring data science expertise.

Decomposition Trees: This AI visual analyzes multidimensional data, using statistical heuristics to automatically identify and drill down into dimensions with maximum variance for interactive root-cause analysis.

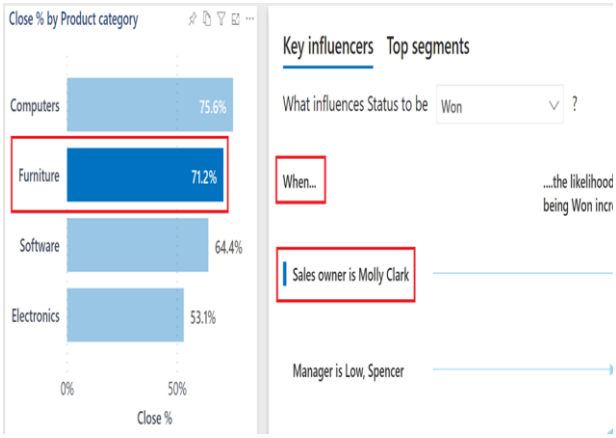


Figure 8: Key Influencers

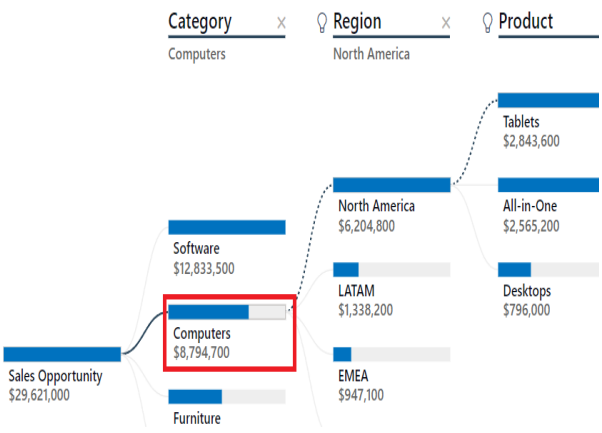


Figure 9: Decomposition Tree

XAI Feature	Core Analytical Functionality	Underlying AI/XAI Technology	Primary Business Application
Explain Data	Analyzes a single data point to explain variance and highlight anomalies	Additive Feature Attribution (e.g., SHAP, LIME principles)	Understanding anomalous metrics, immediate root cause analysis
Key Influencers	Ranks features driving a specific categorical or continuous outcome	Regression, Decision Trees, Feature Importance Scoring	Identifying customer churn drivers, optimizing sales conversion factors
Decomposition Tree	AI-guided drill-down path selection based on statistical variance	Statistical heuristics, Variance maximization algorithms	Ad-hoc exploration, deep dive into complex hierarchical data structures

International Research Journal of Innovative Engineering (IRJIE)

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 8, Issue 4, July- August 2024|

DOI:10.15680/IRJIE.2024.0804004

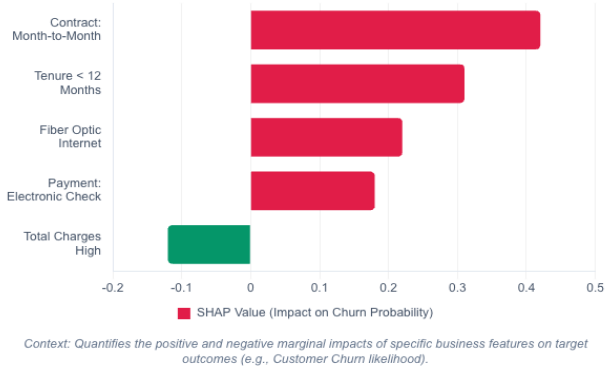


Figure 10: Explainable Artificial Intelligence (XAI): Model Interpretability via SHAP Values and Key Influencer Visuals

VIII. ECONOMIC IMPACT AND ENTERPRISE VALUE REALIZATION

The ultimate measure of any technological advancement in enterprise systems is its tangible impact on operational efficiency, profitability, and measurable cost savings. The deep integration of AI into BI platforms yields a compounding return on investment (ROI) by fundamentally altering how an organization consumes and acts upon data.

A. Quantifying the ROI of AI-Powered BI

Deploying intelligent BI tools speeds time-to-value for enterprise data initiatives. Analyses show that extensive AI-powered BI deployments (such as Power BI within Microsoft 365 E5) can yield over 321% ROI with a payback period under six months.

Key factors driving this economic impact include:

- **Productivity Savings:** NLP and generative tools enable non-technical users to perform self-service data exploration. This reduces reliance on specialized BI teams, saving thousands of labor hours and eliminating reporting bottlenecks.
- **Improved Profitability:** AI instantly flags market shifts, churn drivers, supply bottlenecks, or fraud, allowing managers to optimize operations proactively.
- **Infrastructure Optimization:** Consolidating legacy tools into a cloud-native platform with Medallion architecture lowers ownership costs. Cloud elasticity limits compute resource consumption to active model training or query execution, saving infrastructure costs.

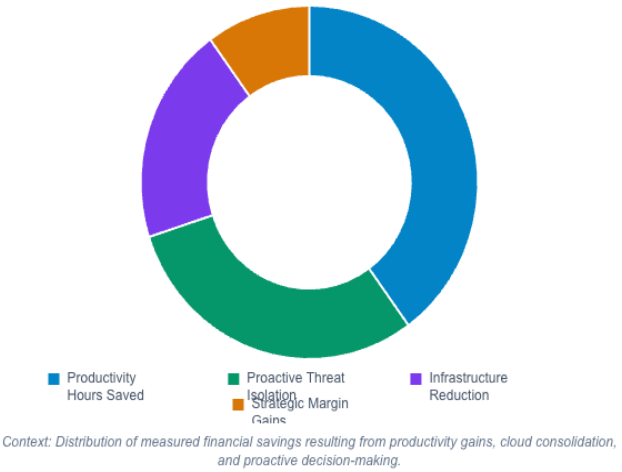


Figure 11: Enterprise Financial Value Driver Breakdown

International Research Journal of Innovative Engineering (IRJIE)

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 8, Issue 4, July- August 2024|

DOI:10.15680/IRJIE.2024.0804004

B. Strategic and Cultural Benefits

AI-powered BI offers profound strategic benefits beyond cost savings. Dashboards evolve into proactive assistants that explain *why* changes occurred rather than just *what* changed. Reusing models from the Silver and Gold data layers connects advanced data science directly to business units through the BI presentation layer. Ultimately, this builds a data-driven culture with decisions grounded in machine-validated evidence.

This study evaluates how enterprise BI combines cloud infrastructure, AutoML, cross-domain semantic parsing, Isolation Forests, and Explainable AI to transform static reports into forward-looking insights.



Figure 12: Big Data Analytics & Artificial Intelligence for Creating Data-Driven Value

IX. CONCLUSION

Combining cloud-native big data platforms, BI, and AI shifts analytics from historical reporting to intelligent prescription. Built on research spanning early BI evolution to Explainable AI, modern BI makes enterprise-grade machine learning accessible.

Medallion architectures orchestrate data flows from ingestion to deployment. Abstracting algorithms with AutoML, NLP, and semantic parsing democratizes conversational data access. Models like Isolation Forest for real-time anomaly detection and SHAP for interpretability ensure accurate, robust, and transparent systems.

AI is now the core engine of modern business strategy. Adopting AI-powered BI capabilities drives operational efficiency, uncovers market dynamics, and delivers a strong competitive advantage.

REFERENCES

1. Business Intelligence and Analytics: From Big Data to Big Impact, https://misq.umn.edu/misq/article/36/4/1165/1483/Business-Intelligence-and-Analytics-From-Big-Data
2. Competing on Analytics - Tom Davenport, https://www.tomdavenport.com/book/competing-on-analytics/
3. Competing on Analytics: The New Science of Winning - Google Books, https://books.google.com/books/about/Competing_on_Analytics.html?id=n7Gp7Q84hcsC
4. Competing on Analytics by Thomas H. Davenport | Goodreads, https://www.goodreads.com/book/show/414001.Competing_on_Analytics
5. Competing on Analytics, https://nucleusanalytics.com/blog/harvard-business-review-competing-on-analytics/
6. Automated Machine Learning: Methods, Systems, Challenges (The, https://northwindbook.com/book/9783030053178
7. Automated Machine Learning: Methods, Systems, Challenges (The, https://www.abebooks.com/9783030053178/Automated-Machine-Learning-Methods-Systems-3030053172/plp
8. Automated machine learning : Methods, systems, challenges, https://www.aiu.edu.sy/public/en/electronic-book/automated-machine-learning-methods-systems-challenges
9. Natural language interfaces to databases – an introduction, https://www.cambridge.org/core/journals/natural-language-engineering/article/natural-language-interfaces-to-databases-an-introduction/21C30448C70DD4988E6DA0D54205FB56
10. Spider: A Large-Scale Human-Labeled Dataset for Complex ... - arXiv, https://arxiv.org/html/1809.08887v5
11. A Comparative Analysis of Machine Learning Models as Key GDP, http://ieeexplore.ieee.org/iel8/6287639/10820123/10938555.pdf
12. Full article: Forecasting interrupted time series - Taylor & Francis, https://www.tandfonline.com/doi/full/10.1080/01605682.2024.2395315
13. Forecasting Multivariate Time Series: A Comparison of Machine, https://www.mdpi.com/2571-9394/8/4/73
14. Isolation forest - Wikipedia, https://en.wikipedia.org/wiki/Isolation_forest

| ISSN: 2395-0560 | www.irjie.org |editor@irjie.org| A Bimonthly, Peer-Reviewed, Scholarly Journal|

|Volume 8, Issue 4, July- August 2024|

DOI:10.15680/IRJIE.2024.0804004

15. Isolation Forest - IEEE Computer Society, https://www.computer.org/csdl/proceedings-article/icdm/2008/3502a413/12OmNqFrGue
16. Anomaly Detection with Isolation Forest - MATLAB & Simulink, https://www.mathworks.com/help/stats/anomaly-detection-with-isolation-forest.html
17. A Unified Approach to Interpreting Model Predictions - NIPS, https://proceedings.neurips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf